THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 6, pp. 3625–3632, February 5, 2010

Catalytic Mechanism of Human α-Galactosidase*[S]

Received for publication, August 26, 2009, and in revised form, November 6, 2009 Published, JBC Papers in Press, November 25, 2009, DOI 10.1074/jbc.M109.060145

Abigail I. Guce[‡1], Nathaniel E. Clark[§1], Eric N. Salgado[§], Dina R. Ivanen[¶], Anna A. Kulminskaya[¶], Harry Brumer III[‖], and Scott C. Garman[‡§2]

From the Departments of [§]Biochemistry and Molecular Biology and [‡]Chemistry, University of Massachusetts, Amherst, Massachusetts 01003, the [¶]Molecular and Radiation Biophysics Division, Petersburg Nuclear Physics Institute, Russian Academy of Science, Orlova Roscha, Gatchina 188300, Leningrad District, Russia, and the [‖]Department of Biotechnology, Royal Insitute of Technology (KTH), 10691 Stockholm, Sweden

The enzyme α-galactosidase (α-GAL, also known as α-GAL A; E.C. 3.2.1.22) is responsible for the breakdown of α-galactosides in the lysosome. Defects in human α-GAL lead to the development of Fabry disease, a lysosomal storage disorder characterized by the buildup of α-galactosylated substrates in the tissues. α-GAL is an active target of clinical research: there are currently two treatment options for Fabry disease, recombinant enzyme replacement therapy (approved in the United States in 2003) and pharmacological chaperone therapy (currently in clinical trials). Previously, we have reported the structure of human α-GAL, which revealed the overall structure of the enzyme and established the locations of hundreds of mutations that lead to the development of Fabry disease. Here, we describe the catalytic mechanism of the enzyme derived from x-ray crystal structures of each of the four stages of the double displacement reaction mechanism. Use of a difluoro-α-galactopyranoside allowed trapping of a covalent intermediate. The ensemble of structures reveals distortion of the ligand into a $^{1}S_{3}$ skew (or twist) boat conformation in the middle of the reaction cycle. The high resolution structures of each step in the catalytic cycle will allow for improved drug design efforts on α-GAL and other glycoside hydrolase family 27 enzymes by developing ligands that specifically target different states of the catalytic cycle. Additionally, the structures revealed a second ligand-binding site suitable for targeting by novel pharmacological chaperones.

The degradation of macromolecules, including glycopeptides and glycolipids, occurs in the lysosome via catabolic enzymes. For example, glycosidases cleave the oligosaccharides from glycoproteins and glycolipids into smaller components used by the cell. One such lysosomal enzyme is α-galactosidase (α-GAL,[3] also known as α-GAL A; E.C. 3.2.1.22), which catalyzes the removal of a terminal α-galactose residue from polysaccharides, glycolipids, and glycopeptides (1). α-GAL also has the ability to convert human blood group B to blood group O (2, 3). The reaction catalyzed by α-GAL is shown in Fig. 1*A*.

Defects in human α-GAL lead to Fabry disease, an X-linked inherited disorder affecting 1 in every 40,000 males characterized by chronic pain, vascular degeneration, cardiac abnormalities, and other symptoms (1). The disease displays distinct phenotypes correlated with the amount of residual enzymatic activity: a severe form affecting multiple organ systems including the eyes, liver, kidney, and heart; and a milder phenotype with symptoms restricted to cardiac and/or renal abnormalities. The severe phenotype generally results from a complete loss of enzymatic activity in affected individuals, whereas patients with milder phenotypes typically show some residual enzyme activity. The data base of independent Fabry disease mutations now numbers in the hundreds, from thousands of patients. Most Fabry disease patients have a single point mutation in their *GLA* gene coding for α-GAL, and over 400 missense and nonsense mutations have been identified in different patients (4). Fabry disease is a lysosomal storage disorder, a family of over 40 diseases characterized by the accumulation of a substrate in the absence of a functional lysosomal enzyme. The family includes inherited diseases such as Tay-Sachs, Sandhoff, and Gaucher diseases.

Lysosomal storage diseases are active subjects of clinical research. In Gaucher disease (a defect in the lysosomal enzyme β-glucosidase), enzyme replacement therapy in pediatric patients successfully treats disease and is currently in use in over 3500 patients from 55 countries (5). Gaucher disease was the first lysosomal storage disease approved for treatment with recombinant enzyme replacement therapy in 1994 (6), and Fabry disease became the second in 2003 (7, 8). Because recombinant enzymes typically will not cross the blood-brain barrier, enzyme replacement therapy is unsuitable for patients with lysosomal storage disorders that manifest neurological symptoms (9), and pharmacological chaperone therapy has been proposed instead. Fabry disease was the first lysosomal storage disease to show the potential efficacy of pharmacological chaperone therapy, with two small molecule compounds, 1-deoxygalactonojirimycin (DGJ) (10, 11) and galactose (12), proposed as pharmacological chaperones in Fabry patients.

* This work was supported, in whole or in part, by National Institutes of Health Grant R01 DK76877 (to S. C. G.) and by National Science Foundation Integrative Graduate Education and Research Traineeship 0654128 (to N. E. C.).

The atomic coordinates and structure factors (codes 3HG2, 3HG3, 3HG4, and 3HG5) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Movie S1.

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed: Dept of Biochemistry & Molecular Biology, 710 North Pleasant St., University of Massachusetts, Amherst, MA 01003. Fax: 413-545-3291; E-mail: garman@biochem.umass.edu.

[3] The abbreviations used are: α-GAL, α-galactosidase; DGJ, 1-deoxygalactonojirimycin; GH, glycoside hydrolase; PEG, polyethylene glycol; PDB, Protein Data Bank; TNP-2,2-di-F-α-Gal, 2′,4′,6′-trinitrophenyl-2-deoxy-2,2-difluoro-α-D-galactopyranoside.

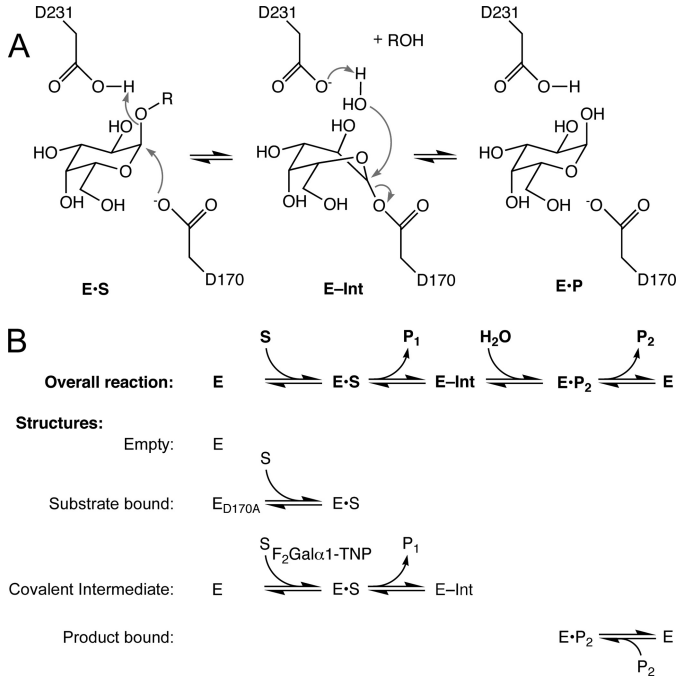

FIGURE 1. **α-GAL reaction and trapping stages for crystallographic analysis.** *A*, double displacement reaction mechanism in human α-GAL. Asp-170 acts as the nucleophile, and Asp-231 acts as an acid and then a base over the course of the reaction cycle. *B*, structures of the different stages in the catalytic cycle. Empty enzyme (*blue*) required a cryoprotectant sterically excluded from the active site. The substrate-bound structure (*green*) resulted from deletion of the active site nucleophile, followed by addition of a disaccharide substrate. The covalent intermediate (*yellow*) used a difluoro-substituted galactoside to slow the second stage of the reaction. The product-bound structure (*red*) resulted from product inhibition of the enzyme. The color scheme is maintained throughout.

α-GAL is an α-retaining enzyme, where both substrate and product have anomeric carbons with α configurations. The mechanism of α-retaining glycosidases was originally proposed by Koshland to be a double displacement reaction mechanism, where two consecutive nucleophilic attacks on the anomeric carbon lead to overall retention of the anomeric configuration (13, 14). This reaction mechanism requires two carboxylates, one acting as nucleophile and one acting as an acid/base (15). The mechanism is proposed to go through oxocarbenium ion-like transition states at two stages in the cycle. In glycoside hydrolases, addition of electronegative fluorine atoms to the 2 or 5 position of the hexoside slows the reaction, allowing for trapping of a covalent intermediate (16). This strategy, combined with mutation of the acid/base carboxylate, helped clarify the mechanism of lysozyme (17) after many years of study. In glycoside hydrolase family 27 (GH27), which includes human α-GAL and 395 other sequences (18), fluorinated substrates have been used to map the catalytic nucleophile by mass spectrometry (19, 20). The identical catalytic mechanism is conserved across glycoside hydrolase clan D (21), which includes 1927 sequences in families 27, 31, and 36 (18).

Previously, we reported structures of human α-GAL at 3.25 and 3.45 Å resolution (22). Additionally, others have recently determined three structures of partially deglycosylated human α-GAL at resolutions of 2.2–3.0 Å (23). Taken as a group, the five structures showed no significant changes in the protein upon binding of ligands including galactose or the pharmacological chaperone DGJ. We used a combination of genetic engineering, microseeding techniques, and synthetic substrates to improve the resolution of the crystals to 1.9 Å and to interrogate the catalytic mechanism of the enzyme. Here we report a series of four new structures of the human α-GAL glycoprotein representing each stage in the catalytic cycle (Fig. 1*B*). The covalent intermediate stage shows a novel conformation of the ligand in the active site, a distorted skew boat conformation of the sugar. The structures reveal conformational changes in the ligand over the course of the reaction cycle, suggesting novel mechanism-based inhibitors. Additionally, the structures reveal a second sugar-binding site distal from the active site, which is ideal for targeting by a new class of pharmacological chaperones.

EXPERIMENTAL PROCEDURES

Molecular Biology

Human α-GAL was expressed in *Trichoplusia ni* (Tn5) insect cells. The wild-type *GLA* gene was subcloned by PCR from a pOTB7 vector (OpenBiosystems) using Phusion DNA Polymerase (Finnzymes) with forward (5′-ACA ATG CAG CTG AGG AAC CCA GAA CTA CAT-3′) and reverse (5′-TTA ATG ATG ATG ATG ATG ATG AAG TAA GTC TTT TAA TGA CAT-3′) primers. The purified PCR product was incubated at 72 °C for 15 min with Taq polymerase (New England Biolabs) to add single 3′-deoxyadenosine overhangs for subsequent TOPO TA cloning into a pIB/V5-His-TOPO vector (Invitrogen). The resulting pIB/V5-His-TOPO-α-GAL construct contains the wild-type α-GAL sequence (including the N-terminal signal sequence) and a C-terminal His_6 tag.

The D170A mutant was generated from the above wild-type construct by site-directed mutagenesis using forward (5′-[PO_4]CTG CTA AAA TTT **GCT** GGT TGT TAC TGT GAC AG-3′, mutation in bold) and reverse (5′-[PO_4]ATC TAC TCC CCA GTC AGC AAA GGT CTG-3′) PCR primers and confirmed by sequencing.

Cell Transfection

Approximately 2.0×10^6 adherent Tn5 cells were transfected with a 2-ml mixture containing SFX-Insect medium (Hyclone), 1.8 μg of plasmid DNA, and Hilymax (Dojindo Labs). After 3 days, 100 μg/ml blasticidin in fresh SFX was added to select for stable transfectants, and supernatants were tested for protein expression. Stable adherent cells were resuspended in SFX medium for larger scale suspension cultures.

Protein Expression and Purification

Stable cells expressing D170A α-GAL mutant protein were grown to $5–6 \times 10^6$ cell/ml for 3 days. The supernatant was concentrated and buffer exchanged with nickel binding buffer (50 mM Na_3PO_4, 500 mM NaCl, and 20 mM imidazole, pH 7.0) using a Prep/Scale tangential flow filtration cartridge (Millipore). Retentate was loaded onto a Ni^{2+}-Sepharose 6 Fast Flow column (GE Healthcare) and eluted with a gradient of 0–50% elution buffer (50 mM Na_3PO_4, 500 mM NaCl, and 1 M imidazole, pH 7.0). Fractions containing pure α-GAL, as determined by SDS-PAGE analysis, were pooled, concentrated to 2.0 mg/ml

using a 10-kDa molecular mass cutoff spin concentrator (Sartorius), and stored in 10 mM Na_3PO_4 buffer, pH 6.5.

Kinetic Assays

α-GAL hydrolysis of the synthetic substrate *para*-nitrophenyl-α-galactose (pNP-α-Gal) (Toronto Research Chemicals) at 37 °C was monitored by absorbance at 400 nm and an extinction coefficient of 18.1 mM^{-1} cm^{-1}. 1 μg of enzyme in 100 mM citrate/phosphate buffer pH 4.5 was added to 8 substrate concentrations of pNP-α-Gal from 0.1 to 50 mM. Every minute for 10 min, sample absorbance was measured after adding 200 mM Na_3BO_3 buffer, pH 9.8. Statistics were calculated from triplicate measurements. K_m, V_{max}, and k_{cat} were calculated in KaleidaGraph from a weighted fit of Michaelis-Menten hyperbola.

Sugars and Sugar Analogues

TNP-2,2-di-F-α-Gal was prepared as described (14, 19). Thioethyl-α-galactose was prepared from 2,3,4,6-tetraacetylthioethyl-α-D-galactose (Toronto Research Chemicals) by treatment with HCl for 1 h and neutralization with NaOH. Mass spectrometry confirmed that the hydrolysis of the four acetyls had run to completion. Galactose, lactose, and melibiose were obtained from Sigma.

Crystallization and X-ray Data Collection

Empty, Intermediate, and Product—Crystals were grown as described in Ref. 22, except that microseeding techniques increased the size of the crystals from 100 to 400 μm in the longest dimension, leading to much improved diffraction. Seed crystals were grown in 25% PEG 4000, 200 mM $(NH_4)_2SO_4$, and 100 mM $NaCH_3COO$, pH 4.6 mixed 1:1 with 9 mg/ml α-GAL stock. The crystals were crushed, diluted 1:100,000, and added to a 1:1:1 mixture of crystallization buffer, protein stock, and water. Crystals appeared in the seeded solution after a few days. Crystals were harvested into buffer containing 30% PEG 4000, 200 mM $(NH_4)_2SO_4$, and 100 mM $NaCH_3COO$. Crystals were then transferred to similar solutions supplemented with ligand: 15% α-lactose (empty structure), 20 mM TNP-2,2-di-F-α-Gal (covalent intermediate), or 5% D-(+)-galactose (product-bound). For the covalent intermediate structure, salt concentrations were reduced to 100 mM $(NH_4)_2SO_4$ and 50 mM $NaCH_3COO$, and 15% *N*-acetylglucosamine was used as a cryoprotectant. Five crystals were transferred into the TNP-2,2-di-F-α-Gal solution; all but one shattered immediately after transfer. Ice rings limited the quality of the galactose-soaked crystal diffraction data. After flash-cooling the crystals, diffraction data were collected on a RU-H3R generator (Rigaku) and processed with HKL2000 (24). The $P3_221$ crystals diffract anisotropically, with ~1.9 Å resolution limits along the *c* axis and ~2.8 Å diffraction limits in perpendicular directions; we limited the overall resolution to 2.3 Å based upon I/σ_I criteria. The isotropic integration of the anisotropic diffraction data resulted in poor merging statistics. The structures were phased by Fourier synthesis using the 3.25 Å human α-GAL coordinates (PDB code 1R46) (22).

Substrate—The crystals of Asp-170 α-GAL mutant were obtained by hanging-drop vapor diffusion at 20 °C. Crystals were obtained from a 1:1 mixture of reservoir solution (12% PEG 8K, 0.1 M sodium cacodylate pH 6.5, and 22 mM $Mg(CH_3COO)_2$) and 2.0 mg/ml protein in 10 mM Na_3PO_4, pH 6.5. Prior to freezing and data collection, crystals were transferred stepwise into reservoir solution with 100 mM melibiose (Sigma), and then into a cryoprotectant solution (40% PEG 8K, 0.1 M sodium cacodylate pH 6.5, 22 mM $Mg(CH_3COO)_2$, 20% ethylene glycol, and 100 mM melibiose). Crystals were flash-cooled in liquid nitrogen, and x-ray data were collected at 100 K at beamline X25 at the Brookhaven National Laboratory. X-ray images were processed using HKL2000 (24) in space group $P2_12_12_1$ and phased by molecular replacement in AMoRe (25) using the coordinates of the galactose-soaked crystal.

For all structures, atomic models were built using the program O (26), with refinement in REFMAC5 (25). Ramachandran plots were computed using PROCHECK (27). Coordinates were superimposed using LSQMAN (28), and accessible surface areas were calculated in AREAIMOL (25). Molecular docking studies were performed in O (26) and in Pymol (29). Figures were made in MolScript (30) and POVScript+ (31). Coordinates and structure factors are deposited in the Protein Data Bank under accession codes 3HG2, 3HG3, 3HG4, and 3HG5.

RESULTS

Overall Description of the Structures

Human α-GAL is a homodimer where each monomer has two domains, an N-terminal $(\beta/\alpha)_8$ barrel containing the active site and a C-terminal antiparallel β domain. The homodimer contains two active sites separated by 43 Å. Previous reports have suggested cooperativity between the two active sites (with a Hill coefficient of 1.9) (32), but the three-dimensional structures show no evidence for cooperativity within the dimer: there are no substantial differences between the two monomers in the asymmetric unit (root mean square deviation of 0.4–0.5 Å for 390 Cα atoms), or among the four steps in the catalytic cycle.

Stages in the Catalytic Cycle

We determined four crystal structures representing each stage in the catalytic cycle of human α-GAL: the empty enzyme at 2.3 Å, the substrate-bound enzyme at 1.9 Å, the covalent intermediate at 2.3 Å, and the product-soaked enzyme at 2.3 Å. Overall the four structures give a snapshot of the catalytic mechanism of the enzyme at each stage. Crystallographic statistics are shown in Table 1. Omit map electron density of the active site contents and interaction maps of the active sites show that most protein:ligand interactions are conserved during the cycle (Fig. 2). Details of each stage in the cycle follow below.

Empty Active Site—Determining the structure of the empty enzyme was not straightforward. Using 20% glycerol or ethylene glycol as a cryoprotectant resulted in an active site containing bound cryoprotectant. The use of larger cryoprotectants (glucose and lactose) that are sterically occluded from the active site allowed for determination of the structure of the empty enzyme.

The empty enzyme shows a collection of water molecules that approximate the locations of the atoms in the galactose ligand (Fig. 2*A*). Three waters reflect the positions of oxygens

TABLE 1
Crystallographic statistics

Active site:	Empty	Substrate	Intermediate	Product
PDB code:	**3HG2**	**3HG3**	**3HG4**	**3HG5**
Protein				
Protein sequence	Wild type	D170A	Wild type	Wild type
Space group	$P3_221$	$P2_12_12_1$	$P3_221$	$P3_221$
Cell lengths, Å	90.8, 90.8, 217.2	59.5, 106.1, 181.7	90.2, 90.2, 216.6	90.8, 90.8, 217.2
Cell angles, °	90, 90, 120	90, 90, 90	90, 90, 120	90, 90, 120
X-ray data				
X-ray source	Cu anode	BNL X25	Cu anode	Cu anode
Wavelength, Å	1.54	1.0085	1.54	1.54
Resolution (last shell)	50-2.3 (2.38-2.3)	50-1.90 (1.97-1.9)	50-2.3 (2.38-2.3)	50-2.3 (2.38-2.3)
Observations	637,404	1,248,391	739,029	398,547
Unique observations	45,612	90,308	46,337	47,076
Completeness, % (last shell)	97.2 (95.4)	98.6 (88.4)	99.7 (100.0)	99.9 (100.0)
Multiplicity (last shell)	14.0 (14.2)	13.8 (9.9)	15.9 (16.0)	8.5 (8.4)
R_{sym} (last shell)[a]	0.115 (0.818)	0.088 (0.755)	0.112 (0.944)	0.099 (0.989)
I/σ_I (last shell)	25.5 (3.5)	30.2 (3.1)	28.3 (3.4)	21.4 (2.3)
Refinement				
R_{work}/R_{free},[b] %	17.6/20.2	16.5/19.7	16.5/22.1	19.1/22.7
No. of atoms	6767	7539	7251	6694
Protein	6255	6377	6262	6255
Carbohydrate	156	216	276	167
Water	329	893	682	236
Other	27	53	31	36
Average B, Å^2	36.4	23.1	49.8	43.4
Ramachandran Plot[c]				
Favored (%)	90.7	92.1	90.0	90.5
Allowed (%)	8.4	7.0	9.4	8.4
Generous (%)	0.6	0.9	0.4	0.7
Forbidden (%)	0.3	0.0	0.1	0.3
RMS deviations				
Bonds (Å)	0.008	0.006	0.008	0.008
Angles (°)	1.128	1.058	1.159	1.175

[a] $R_{sym} = \Sigma_h\Sigma_i|I_{h,i} - <I_h>|/\Sigma_h\Sigma_i|I_{h,i}|$, where $I_{h,i}$ is the i^{th} intensity measurement of reflection h and $<I_h>$ is the average intensity of that reflection.
[b] R_{work}, $R_{free} = \Sigma_h||F_P| - |F_C||/\Sigma_h|F_P|$, where F_C is the calculated and F_P is the observed structure factor amplitude of reflection h for the working or free set, respectively.
[c] Ramachandran statistics are calculated in PROCHECK.

O2 (0.92 Å away), O3 (0.27 Å), and O6 (0.21 Å) of the galactose, and two waters reflect the positions of carbons C1 (0.51 Å away) and C4 (1.1 Å) of the galactose.

Substrate-bound—To determine the structure of the substrate-bound crystal, we initially soaked a non-hydrolyzable substrate analogue into the active site. In two different experiments, 25 mM thioethyl-α-D-galactose was added to the crystals; however the active site was unoccupied in each of those experiments (data not shown). We modeled the thioethyl-α-D-galactose structure based upon the thiomethyl-α-D-galactosyl component of the antibiotic clindamycin (33), and the thioethyl group had steric clashes with the protein residues in their normal location.

We then turned to engineering the protein to capture substrate-bound α-GAL. We made a mutant protein D170A, lacking the active site nucleophile, in insect cells. We performed kinetic studies of wild-type and D170A mutant enzymes using the synthetic substrate *para*-nitrophenyl-α-D-galactoside. The wild-type enzyme had a K_m of 8.3 ± 0.5 mM and a k_{cat} of 63.5 ± 0.1 s^{-1}, (comparable to the two recombinant α-GALs used in enzyme replacement therapy, Ref. 34) whereas the mutant enzyme had no detectable activity ($k_{cat} < 0.1$ s^{-1}). We grew crystals of the D170A mutant protein and soaked 100 mM melibiose (D-Gal-α1–6-D-Glc) into the crystals. The electron density reveals that the melibiose substrate binds to the active site (Fig. 2*B*) with the glucose portion of the substrate extending out of the active site of the enzyme. The galactoside portion of the melibiose substrate is in standard $^4C_1$ chair conformation. As expected for an enzyme with exquisite specificity for the terminal α-galactoside but little specificity beyond the glycosidic linkage, the enzyme makes specific interactions with each functional group on the galactoside portion of the melibiose, but few interactions with the glucoside portion of the disaccharide. Interactions between the galactoside of melibiose and the protein are summarized in Fig. 2*B*.

Covalent Intermediate—To trap a covalent intermediate, we used the synthetic substrate TNP-2,2-di-F-α-Gal (14, 19). The fluoro substituents attached to the 2 position of the galactopyranose ring destabilize the two oxocarbenium ion-like transition states in the catalytic mechanism, but the excellent leaving group TNP allows the first nucleophilic attack to proceed. The second nucleophilic attack by a deprotonated water molecule is slowed considerably, allowing us to trap the intermediate (Fig. 3*A*). Cryogenic temperatures increased the lifetime of the covalent complex, allowing collection of complete x-ray data. Although fluoro sugars have been used in mass spectrometry experiments to identify the catalytic nucleophile of α-GAL from *Phanerochaete chrysosporium* (19) and *Coffea arabica* (20), this is the first three-dimensional structure of a covalent complex in the family. The trapped fluoro intermediate is likely to be identical in conformation to the natural intermediate, by extension from studies on *Drosophila melanogaster* α-mannosidase II, where three covalent intermediates show the same conformation independent of fluorine substitution (35).

The electron density for the covalent intermediate shows clear connectivity between C1 of the ligand and Oδ2 of the

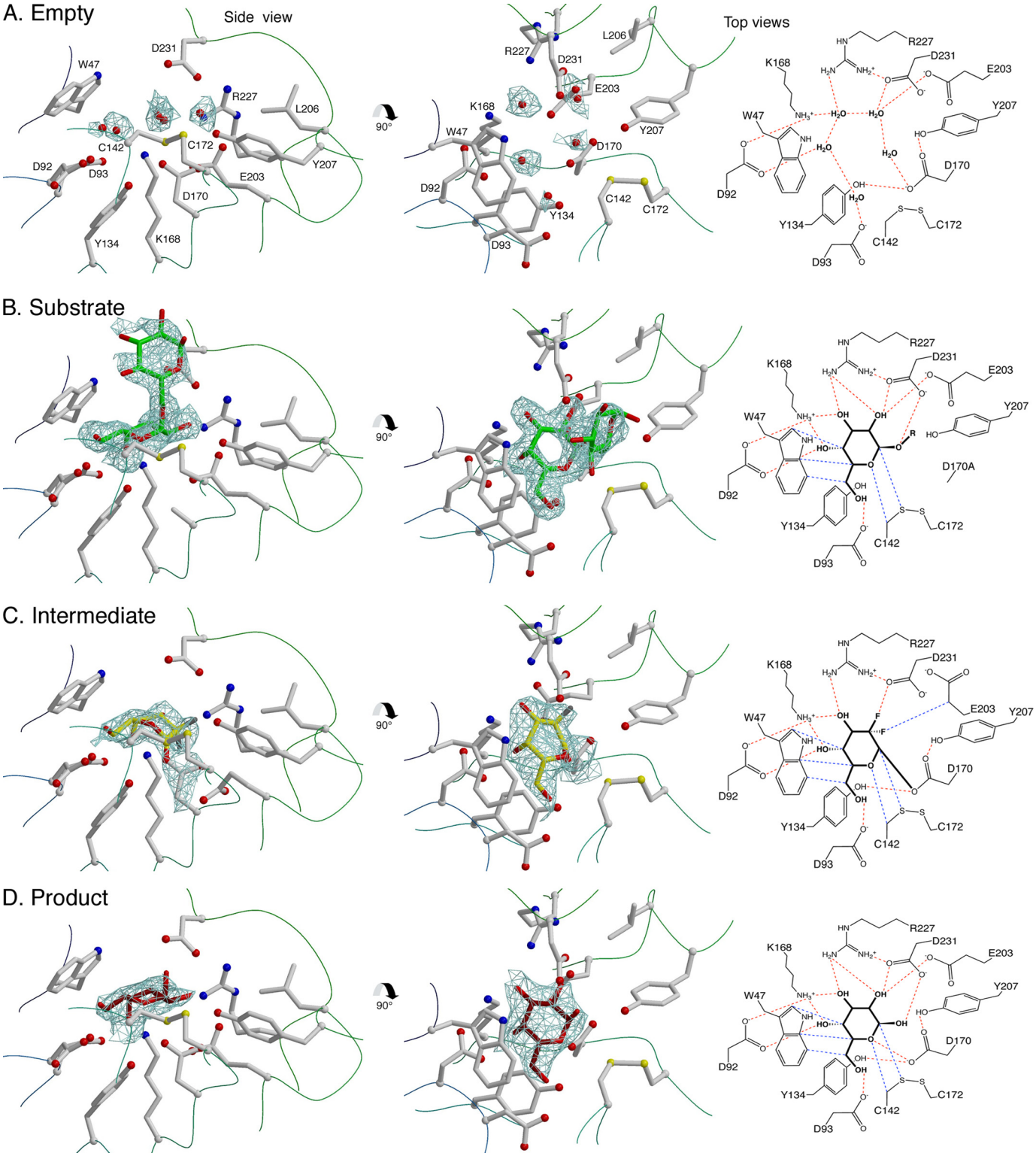

FIGURE 2. **Ligand density and interactions.** *A–D*, *first* and *second columns* show side and top views of the electron density for the ligand in the four different structures. The *third column* shows the interactions around the ligand in the active site, where *red lines* represent hydrogen bonds and *blue lines* represent van der Waals interactions. Empty (*A*), substrate-bound (*B*), covalent intermediate (*C*), and product-bound structures (*D*) are shown, respectively. The electron density corresponds to a σ_A-weighted $2F_o$-F_c total omit map calculated in SFCHECK (25), contoured at 1.5σ in *A* and 2.0σ in *B–D*, with a cover radius drawn around residues and/or waters in the active site.

catalytic nucleophile Asp-170 (Fig. 2*C*). The covalent intermediate is in a $^1S_3$ skew (or twist) boat conformation. The conformation of the covalent intermediate is distorted from the favored chair conformation of the sugar ring. The ligand is held in this distorted conformation at one end by the covalent bond between C1 of the ligand and Oδ2 of Asp-170 and at the other end of the ligand by van der Waals interactions between Trp-47 and the 4-, 5-, and 6 carbons on the β face of the sugar (Fig. 3*B*).

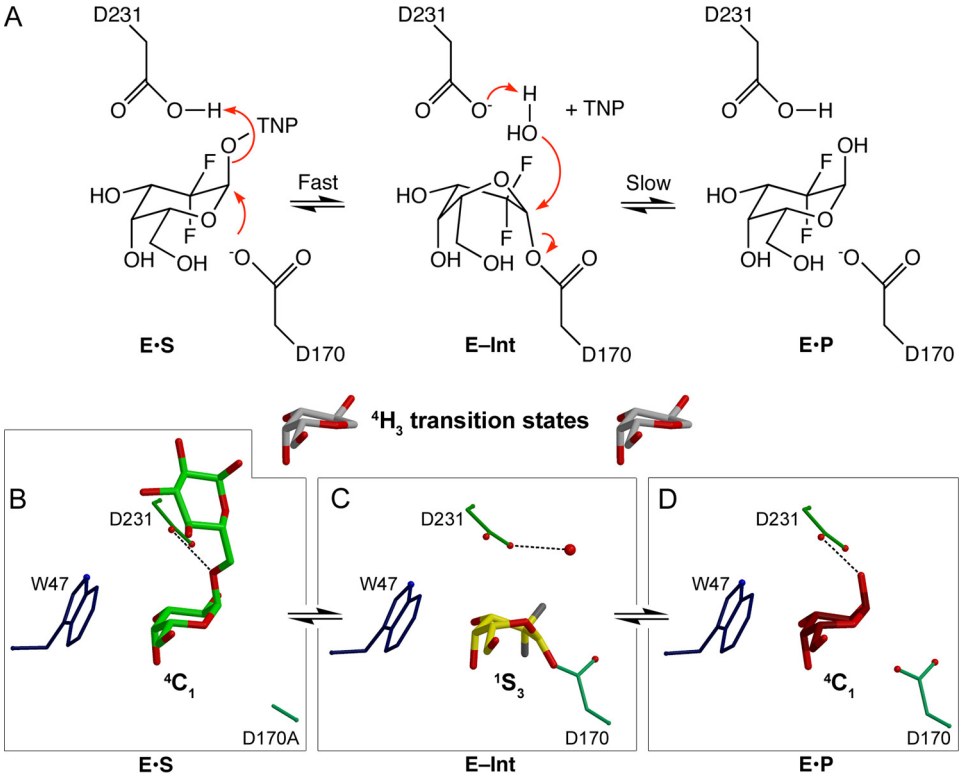

FIGURE 3. **Covalent intermediate and ligand deformation.** *A*, reaction mechanism for the TNP-2,2-di-F-α-Gal substrate. The first step of the reaction remains fast because of the good leaving group, but the second step slows, allowing for trapping of the covalent intermediate species. *B–D*, close-up views of the substrate-bound, covalent intermediate, and product-bound ligands. The catalytic nucleophile Asp-170, the catalytic acid/base Asp-231, and the conserved Trp-47 are labeled. The ligand conformation changes from $^4C_1$ to $^1S_3$ to $^4C_1$ geometry over the course of the reaction. The modeled transition states in $^4H_3$ half-chair conformations are shown in the *insets*.

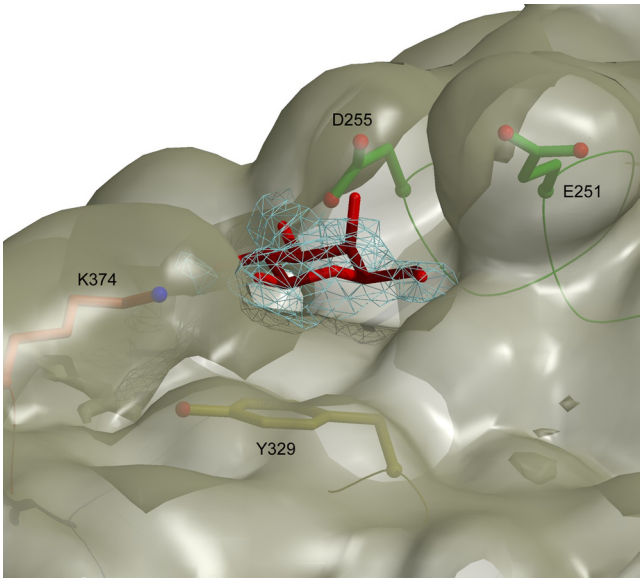

FIGURE 4. **The second ligand-binding site in human α-GAL.** The second ligand-binding site is centered on Tyr-329, at the interface between domain 1 and domain 2 of the structure. A surface drawn around all the atoms reveals a pocket that is selective for β-galactose. The electron density from the galactose-soaked crystal shows a σ_A-weighted $2F_o$-F_c total omit map calculated in SFCHECK, contoured at 1.1σ with a cover radius drawn around the ligand. Residues Glu-251, Asp-255, Tyr-329, and Lys-374 are shown.

Product-bound—We determined the structure of the product-bound crystal by soaking in the catalytic product galactose. Although galactose solution contained a 70:30 mixture of β and α anomers of galactose, the active site selects for the α anomer only (Fig. 2*D*). This selectivity is unique to the animal kingdom enzymes in GH27, as differences in the active site residues in the rice and *Trichoderma reesei* α-GAL structures accommodate both α and β anomers of galactose in the active site (36, 37). The galactose in the product soak is in a standard $^4C_1$ conformation. After the second nucleophilic attack on the anomeric carbon, the ligand reverts to a low energy conformation from its distorted state in the covalent intermediate structure.

A Second Ligand-binding Site in Human α-GAL

In our crystal soaks, we have discovered a second ligand-binding site on the surface of the α-GAL molecule, at the interface between the two domains of the monomer. In three of the four structures presented here, a monosaccharide packs on the surface of Tyr-329 and forms hydrogen bonds with Asp-255 and Lys-374. (In the fourth structure, the D170A melibiose soak, the disaccharide is sterically excluded from the second binding site, and a PEG molecule is found there.)

The second ligand-binding site prefers the β anomer of galactose (Fig. 4). When D-galactose (approximately a 70:30 mixture of β and α anomers) is added to the crystals, the primary binding site (the active site) binds only the α anomer, and the secondary binding site binds only the β anomer. The secondary binding site binds ligand when the crystals are soaked with 15% lactose (empty active site), 100 mM melibiose (substrate bound in active site), or 20% galactose (product in active site). Additionally, when 20% glucose is added to the crystals, glucose binds in the second binding site (data not shown). The second binding site uses the plane of Tyr-329 to pack against the β face of the monosaccharide, a common protein/carbohydrate interaction (38). Electron density in the second site is likely a result of fortuitous binding in the presence of high concentrations of sugars used as cryoprotectants. The second binding site is comparable in size to the active site: there are 154 Å^2 of protein surface area buried when β-galactose binds to the protein at the second site, larger than the 107 Å^2 of protein surface buried when α-galactose binds in the active site.

DISCUSSION

In this report, we describe four structures of human α-GAL with improved resolution, allowing us to examine the catalytic mechanism of the enzyme in detail. The structures here represent the first member of GH27 with crystallographic evidence

for each step in the reaction mechanism. Given the strict conservation in the binding site residues within the GH27 family and the conserved mechanism across clan D (14), the mechanistic studies presented here extend to the entire clan of over 1900 protein sequences. To demonstrate the mechanism, we modeled the entire catalytic cycle of the enzyme, including the conformational changes of the ligand over the course of the reaction (supplemental Movie S1).

Pharmacological Chaperones—The discovery of a second ligand-binding site on the α-GAL molecule leads to another approach for pharmacological chaperoning for the treatment of Fabry disease. Pharmacological chaperones bind to the folded conformation of a protein and help to stabilize it. The current generation of pharmacological chaperones for the treatment of Fabry disease, including galactose and DGJ, are substrate and product analogues that bind to the active site of the folded glycoprotein. Because they bind to the active site, they act as competitive inhibitors as well as chaperones; thus they must depart from the active site before the enzyme can hydrolyze substrate. An alternative approach, as noted by the Petsko group (23), might capitalize upon a pharmacological chaperone-binding site distal to the α-GAL active site, which would stabilize the molecule without competitively inhibiting the enzyme. Here we report such a site. The second binding site has distinct substrate specificity from the active site: when we soaked in a mixture of both the α and the β anomers of galactose, the α anomer bound specifically at the active site and the β anomer bound specifically at the second ligand-binding site. Thus, we predict that a new class of pharmacological chaperones can be directed to bind to the second ligand-binding site.

Ligand Geometry in Retaining Glycosidases—The conformational pathways of α-retaining enzymes represent something of a paradox. For the pathways with known structures of covalent intermediates, some (including GH29, GH31, GH38, and now GH27) show distorted sugars in the covalent intermediate, but others (such as GH13 and GH77) show distorted sugars in the noncovalent substrate complex. For example, human α-*N*-acetylgalactosaminidase (GH27) has a strained $^{1}S_{3}$ intermediate (39), *Thermotoga maritima* α-L-fucosidase (GH29) has a strained $^{3}S_{1}$ intermediate (40), *Escherichia coli* α-glucosidase YicI (GH31) has a strained $^{1}S_{3}$ intermediate (41), and *Drosophila melanogaster* Golgi α-mannosidase II (GH38) has a strained $^{1}S_{5}$ intermediate (35). However, the α-amylase family enzymes *Bacillus circulans* cyclodextrin glucanotransferase (GH13) and *Thermus thermophilus* amylomaltase (GH77) have covalent intermediates in the low energy $^{4}C_{1}$ conformation, but distort the sugars in the noncovalent Michaelis complexes (42, 43).

The results presented here clarify the mechanisms of retaining α-glycosidases. When coupled with the other glycosidases with detailed structural information, some general rules emerge. First, retaining enzymes need to align the *syn* lone pair of the nucleophile to attack the accessible side of the anomeric carbon opposite the leaving group. In the α-retaining enzymes α-galactosidase, α-fucosidase, α-glucosidase, and α-mannosidase, little distortion in the substrate is required to achieve the in-line attack of the nucleophile (35, 40, 41). Second, the same sugar conformations can be used very differently in glycosidase reaction mechanisms. For example, in β-retaining glucosidases (and other β-retaining enzymes), access to in-line attack of the anomeric carbon is achieved by distorting the substrate into $^{1}S_{3}$ geometry, and the subsequent covalent intermediate adopts the low energy $^{4}C_{1}$ conformation (42, 43). This is the exact inverse of the α-GAL and α-glucosidase mechanisms (41), where the substrate has the $^{4}C_{1}$ conformation and the covalent intermediate has $^{1}S_{3}$ geometry. On the Stoddart map of pyranose ring interconversions (44), β-glucosidase and α-galactosidase travel the same path between $^{4}C_{1}$ and $^{1}S_{3}$ conformations, but in opposite directions, reminiscent of the α- and β- retaining mannosidases, which traverse a path between $^{1}S_{5}$ and $^{O}S_{2}$ conformations in opposite directions (35).

Mechanism-based Inhibitors—Galactose and DGJ have shown promise as pharmacological chaperones for treatment of Fabry disease, but they are the catalytic product and a product analogue respectively. With pharmacological chaperones for Fabry disease, *in vitro* inhibition correlates with intracellular α-GAL activity enhancement (10), so tighter binding compounds could be promising candidates for pharmacological chaperone therapy. Imino sugars such as DGJ are generally not good mimics of transition state geometry (45), but transition state or covalent intermediate mimics might now be developed from the human α-GAL structures reported here.

The carbohydrate conformations found in the α-GAL catalytic cycle suggest two new classes of inhibitors for human α-GAL (and, by extension, the entire GH27 family). The first class of inhibitors are analogues of the $^{1}S_{3}$ skew boat covalent intermediate. Whereas there are currently no inhibitors of α-GAL with this configuration, other glycosidases have potent inhibitors that mimic the covalent intermediate, such as the Golgi α-mannosidase inhibitor mannostatin (46).

The second class of novel inhibitors suggested by our α-GAL mechanistic studies are transition state analogues. We modeled the conformation of the sugar ring during the two transition states of the catalytic mechanism as $^{4}H_{3}$ (half-chair) conformations (Fig. 3, *insets*). Thus, a new class of inhibitors for α-GAL (and other related family 27 glycosidases) might adopt a $^{4}H_{3}$ half-chair (or the closely related ^{4}E envelope) conformation of the ring. Transition state mimics have not been developed for α-GAL, however, transition state analogues have been developed as selective inhibitors of mannosidases and other glycosidases (45, 47, 48).

The active site residues of GH27 members are highly conserved, so we predict that the two classes of mechanism-based inhibitors above would be effective for all GH27 members and potentially for the entire 1927 glycosidases in clan D. Given the huge variety of catabolic and anabolic processes that use glycosidases (*e.g.* digestion, lysosomal degradation, ER-associated degradation, and biosynthesis), potent inhibitors of glycosidases are valuable commercial and research products (49). Inhibitors of clan D glycosidases would have broad applications, including, for example, modulating leaf development in plants (50), seed germination in legumes (51), and substrate reduction therapy in galactosemia (52).

In summary, we have discovered a second ligand-binding site on human α-GAL that is suitable for binding by a novel pharmacological chaperone. Second, the structures described herein open the door to designing a new family of mechanism-based inhibitors for the entire family and clan of glycosidases.

Acknowledgments—We thank Matt Metcalf for contributions to the project. Replagal was a gift from T. K. T. We thank Konstantin Shabalin (PNPI, Gatchina) for NMR analysis of TNP-2,2-di-F-α-Gal. We gratefully acknowledge Jean Jankonic, Marc Allaire, and Vivian Stojanoff at the National Synchrotron Light Source X6A beam line, funded by the National Institute of General Medical Sciences, National Institute of Health under agreement GM-0080, and Annie Héroux at PXRR beamline X25, funded by the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy, and from the National Center for Research Resources of the National Institutes of Health.

REFERENCES

1. Desnick, R. J., Ioannou, Y. A., and Eng, C. M. (2001) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3733–3774, 8th Ed., McGraw-Hill, New York
2. Liu, Q. P., Sulzenbacher, G., Yuan, H., Bennett, E. P., Pietz, G., Saunders, K., Spence, J., Nudelman, E., Levery, S. B., White, T., Neveu, J. M., Lane, W. S., Bourne, Y., Olsson, M. L., Henrissat, B., and Clausen, H. (2007) *Nat. Biotechnol.* **25,** 454–464
3. Olsson, M. L., and Clausen, H. (2008) *Br. J. Haematol.* **140,** 3–12
4. Garman, S. C. (2007) *Acta Paediatr. Suppl.* **96,** 6–16
5. Brady, R. O. (2003) *Acta Paediatr. Suppl.* **92,** 19–24
6. Brady, R. O. (2003) *Philos. Trans. R Soc. Lond. B Biol. Sci.* **358,** 915–919
7. Schiffmann, R., Kopp, J. B., Austin, H. A., 3rd, Sabnis, S., Moore, D. F., Weibel, T., Balow, J. E., and Brady, R. O. (2001) *JAMA* **285,** 2743–2749
8. Eng, C. M., Guffon, N., Wilcox, W. R., Germain, D. P., Lee, P., Waldek, S., Caplan, L., Linthorst, G. E., and Desnick, R. J. (2001) *N. Engl. J. Med.* **345,** 9–16
9. Desnick, R. J., and Schuchman, E. H. (2002) *Nat. Rev. Genet.* **3,** 954–966
10. Asano, N., Ishii, S., Kizu, H., Ikeda, K., Yasuda, K., Kato, A., Martin, O. R., and Fan, J. Q. (2000) *Eur. J. Biochem.* **267,** 4179–4186
11. Fan, J. Q., Ishii, S., Asano, N., and Suzuki, Y. (1999) *Nat. Med.* **5,** 112–115
12. Frustaci, A., Chimenti, C., Ricci, R., Natale, L., Russo, M. A., Pieroni, M., Eng, C. M., and Desnick, R. J. (2001) *N. Engl. J. Med.* **345,** 25–32
13. Koshland, D. E. (1953) *Biol. Rev. Cambridge Philos. Soc.* **28,** 416–436
14. Brumer, H., 3rd, Sims, P. F., and Sinnott, M. L. (1999) *Biochem. J.* **339,** 43–53
15. Vocadlo, D. J., and Davies, G. J. (2008) *Curr. Opin. Chem. Biol.* **12,** 539–555
16. Zechel, D. L., and Withers, S. G. (2000) *Acc. Chem. Res.* **33,** 11–18
17. Vocadlo, D. J., Davies, G. J., Laine, R., and Withers, S. G. (2001) *Nature* **412,** 835–838
18. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) *Nucleic Acids Res.* **37,** D233–238
19. Hart, D. O., He, S., Chany, C. J., 2nd, Withers, S. G., Sims, P. F., Sinnott, M. L., and Brumer, H., 3rd. (2000) *Biochemistry* **39,** 9826–9836
20. Ly, H. D., Howard, S., Shum, K., He, S., Zhu, A., and Withers, S. G. (2000) *Carbohydr. Res.* **329,** 539–547
21. Comfort, D. A., Bobrov, K. S., Ivanen, D. R., Shabalin, K. A., Harris, J. M., Kulminskaya, A. A., Brumer, H., and Kelly, R. M. (2007) *Biochemistry* **46,** 3319–3330
22. Garman, S. C., and Garboczi, D. N. (2004) *J. Mol. Biol.* **337,** 319–335
23. Lieberman, R. L., D'Aquino, J. A., Ringe, D., and Petsko, G. A. (2009) *Biochemistry*
24. Otwinowski, Z., and Minor, W. (1997) in *Methods in Enzymology : Macromolecular Crystallography, Part A* (Carter, C. W., and Sweet, R. M., eds) pp. 307–326, Academic Press
25. Collaborative Computational Project (1994) *Acta Crystallogr.* **D50,** 760–763
26. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47,** 110–119
27. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallog.* **26,** 283–291
28. Kleywegt, G. J., and Read, R. J. (1997) *Structure* **5,** 1557–1569
29. DeLano, W. L. (2008) *The PyMOL Molecular Graphics System*, DeLano Scientific LLC, Palo Alto, CA
30. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
31. Fenn, T. D., Ringe, D., and Petsko, G. A. (2003) *J. Appl. Crystallogr.* **36,** 944–947
32. Bishop, D. F., and Desnick, R. J. (1981) *J. Biol. Chem.* **256,** 1307–1316
33. Tu, D., Blaha, G., Moore, P. B., and Steitz, T. A. (2005) *Cell* **121,** 257–270
34. Lee, K., Jin, X., Zhang, K., Copertino, L., Andrews, L., Baker-Malcolm, J., Geagan, L., Qiu, H., Seiger, K., Barngrover, D., McPherson, J. M., and Edmunds, T. (2003) *Glycobiology* **13,** 305–313
35. Numao, S., Kuntz, D. A., Withers, S. G., and Rose, D. R. (2003) *J. Biol. Chem.* **278,** 48074–48083
36. Fujimoto, Z., Kaneko, S., Momma, M., Kobayashi, H., and Mizuno, H. (2003) *J. Biol. Chem.* **278,** 20313–20318
37. Golubev, A. M., Nagem, R. A., Brandão Neto, J. R., Neustroev, K. N., Eneyskaya, E. V., Kulminskaya, A. A., Shabalin, K. A., Savel'ev, A. N., and Polikarpov, I. (2004) *J. Mol. Biol.* **339,** 413–422
38. Weis, W. I., and Drickamer, K. (1996) *Annu. Rev. Biochem.* **65,** 441–473
39. Clark, N. E., and Garman, S. C. (2009) *J. Mol. Biol.* **393,** 435–447
40. Sulzenbacher, G., Bignon, C., Nishimura, T., Tarling, C. A., Withers, S. G., Henrissat, B., and Bourne, Y. (2004) *J. Biol. Chem.* **279,** 13119–13128
41. Lovering, A. L., Lee, S. S., Kim, Y. W., Withers, S. G., and Strynadka, N. C. (2005) *J. Biol. Chem.* **280,** 2105–2115
42. Uitdehaag, J. C., Mosi, R., Kalk, K. H., van der Veen, B. A., Dijkhuizen, L., Withers, S. G., and Dijkstra, B. W. (1999) *Nat. Struct. Biol.* **6,** 432–436
43. Barends, T. R., Bultema, J. B., Kaper, T., van der Maarel, M. J., Dijkhuizen, L., and Dijkstra, B. W. (2007) *J. Biol. Chem.* **282,** 17242–17249
44. Stoddart, J. F. (1971) *Stereochemistry of Carbohydrates*, Wiley-Interscience, New York
45. Wicki, J., Williams, S. J., and Withers, S. G. (2007) *J. Am. Chem. Soc.* **129,** 4530–4531
46. Kawatkar, S. P., Kuntz, D. A., Woods, R. J., Rose, D. R., and Boons, G. J. (2006) *J. Am. Chem. Soc.* **128,** 8310–8319
47. Heck, M. P., Vincent, S. P., Murray, B. W., Bellamy, F., Wong, C. H., and Mioskowski, C. (2004) *J. Am. Chem. Soc.* **126,** 1971–1979
48. Tailford, L. E., Offen, W. A., Smith, N. L., Dumon, C., Morland, C., Gratien, J., Heck, M. P., Stick, R. V., Blériot, Y., Vasella, A., Gilbert, H. J., and Davies, G. J. (2008) *Nat. Chem. Biol.* **4,** 306–312
49. Asano, N. (2003) *Glycobiology* **13,** 93R–104R
50. Chrost, B., Kolukisaoglu, U., Schulz, B., and Krupinska, K. (2007) *Planta* **225,** 311–320
51. Reid, J. S., Edwards, M. E., Gidley, M. J., and Clark, A. H. (1992) *Biochem. Soc. Trans.* **20,** 23–26
52. Holton, J. B., Walter, J. H., and Tyfield, L. A. (2001) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), 8th Ed., pp. 1553–1588, McGraw-Hill, New York